Bayesian Optimal Pure Tone Audiometry with Prior Knowledge

NAN GUAN

KTH Electrical Engineering

Master's Degree Project
Stockholm, Sweden January 2011

XR-EE-SIP 2011:001

Master Thesis
KTH
Sound and Image Process Lab
Royal Institute of Technology (KTH)
SE-10044 STOCKHOLM, Sweden

ACKNOWLEDGEMENT

First of all, I would like to express my gratitude to Professor Arne Leijon for his enormous support during this research work.

I would like to also give my most appreciation to my supervisor Svante Stadler who afforded me most of guidance, suggestions and supports on this research work. Thanks for Associate Professor Bert de Vries and all other teachers who have assisted on my experimental work.

Finally, I am deepest thankful to my wife, parents and friends for their supports on my studies and lives.

ABSTRACT

Pure tone hearing threshold measurement is the most basic and common test for diagnosis of hearing loss and for compensation of the loss with hearing instruments. Pure-tone hearing thresholds are usually assessed using a simple standardized method. By employing an optimal strategy, the thresholds can be determined with the same accuracy as the standard method with much less presentations. With prior knowledge extracted from the Beltone's extra database, which contains over 400,000 audiograms including age and gender information, a more detailed prior knowledge will help improve the optimal strategy and efficiency of PTA process. Meanwhile, a graphical user interface is implementing the method with a more direct way to the users (doctors and patients), which makes the optimal process more accessible and easy to control.

Keywords: Bayesian Method, Pure Tone Audiometry, GMM Method, Matlab GUI.

TABLE OF CONTENT

1. INTRODUCTION

Hearing loss is not a harmless disease that can be left untreated; it has tremendous impact to people's routine life and can affect bother their normal communication with other people and their personal emotion-isolation, depression are both reported to affect mental health of hearing impaired people. Hearing instruments and Audiometry are aiming for hearing impairment fitting and are improved to be effective in compensating such loss.

Pure tone Audiometry (PTA) is the basis of selecting and configuring hearing instruments for certain human being, while the traditional method-ascending method- is spending much time on starting stage and at the risk of start over if unexpected suspension happens. Bayesian optimized PTA is aiming for reducing the time spent on the beginning stage of testing to reduce the presentations needed by the procedure, as outcome of previous presentations can be utilized; a suspended test can also be saved to continue later with this technique.

At present time, the procedure of Bayesian optimal PTA-which includes prior knowledge collecting and re-estimation- is mostly done by computer, as a large amount of complex mathematical calculation is concerned. Obviously the original result of computer calculation is incomprehensible and unusable for Audiometry tester, thus a user interface is needed to interpret the result into intuitive way to be accepted by the tester. Matlab based user interface is selected to accomplish the goal for two basic reasons: efficiency at mathematical model calculating; easy configurable and friendly user interface.

This thesis paper just introduces the implementation of the above content:

1. To implement a GUI to allow a practicing audiologist or doctor to effectively benefit from the Bayesian optimal approach.

2. To extend Svante Stadler's previous work [1] by utilizing additional data from the Beltone database. Stadler used only thresholds data but this work will also make it possible to improve test efficiency by including prior knowledge about the statistical dependence between the tested person's hearing thresholds and age and gender.

1.1. Hearing Impairment over the World

According to one of the recent study, 34 million American citizens suffer from different extents of hearing impairment-roughly 11 percent of the whole population. Over the last generation, the number of hearing impairment people has increased with a rate 1.6 times of American population growth [2]. Figure 1-1 shows the trend of hearing loss population in millions in America with projection through 2050 [2]. While in Europe, about 71 million adults aged from 18 to 80 years have hearing loss greater than 25 dB, which is the definition of hearing impairment recognized by the WHO [3]. Here is some statistics for hearing loss population in some countries [4]:

- Germany: 10.2 million(one in five)
- Italy: 7.2 million(one in six)
- Sweden: 1.3 million(one in ten)

Hearing loss can either be conductive or sensorineural (or called cochlea hearing loss), sometime even both. Such loss could result in low sensitive to different frequency sounds or even weak understanding ability to languages. Three major causes of hearing loss are ageing, exposure to noise and disease, infections or drugs.

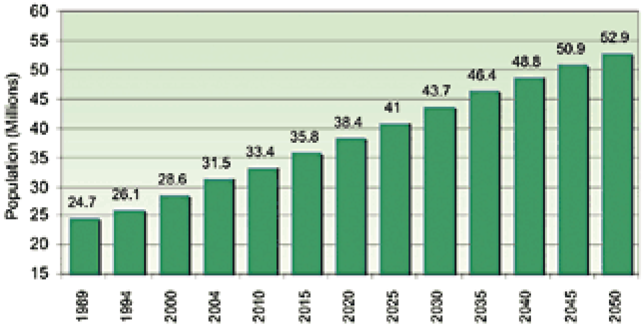

Fig 1-1 Hearing loss population (1989-2004) in America with projections through the year 2050[2]

1.2. Pure Tone Audiometry

Pure tone Audiometry (PTA) is being used worldwide in hearing test for hearing impaired people to identify the hearing loss degree and type, the resulting audiogram is being used as basic information to help doctors diagnose and determine hearing aid treatment. The method relies on patient’s response to pure tone stimuli and only measures the thresholds, which is why it can only be used on adults and children old enough to cooperate with tester. Like most other clinical tests, calibration of the test environment to ISO standard is required before test proceed. During the procedure, PTA uses frequency specific tones to take specific response, and these responses will provide specific configuration of a hearing loss. Although PTA has its clinical advantages, it’s not ideal at identifying all kinds of losses, such as ‘dead regions’.

Currently, there are two standards applied to such identification, international and British standards [5], both based on PTA protocol. The British Society of Audiology is responsible for publishing recommended procedure for PTA.

The standard method for obtaining pure tone hearing thresholds, known as the ascending method, is a procedure which assesses the thresholds of air conduction by going from an inaudible to audible stimulus intensity. Such procedure is evaluated independently at each frequency. There are also other methods like descending method which assesses the thresholds by going from audible to inaudible stimulus intensity, or "up 5-down 10" method which estimates thresholds by reducing 10 dB to positive response and raising 5 dB to negative response.

1.3. Bayesian Optimized Psychoacoustic Measurement

For ascending method, a large amount of time is spent on each frequency to start over from inaudible sound levels, although adjacent thresholds are highly correlated. As a result, it's becoming a time-consuming process which requires a fair amount of presentation data to be collected.

Consequently the most important question turns to be that whether these redundant data collecting can be achieved by a more efficient psychometric procedure or can be accomplished in a more rapid way? If the answer is positive, how many presentations can be avoided, if the procedure of testing is optimized? While adjacent thresholds are taken into account, will the procedure be less redundant? Also, can the procedure be designed so that an optimal threshold estimate can be obtained, even if the experiment is suspended prematurely?

Adaptive procedures have been developed to address this major problem in psychometric measurement [6]. An adaptive experiment procedure is defined by whether the order of new test is determined by the results of previous tests. One feature of adaptive procedures is that knowledge about thresholds will increase as long as the procedure processing. Thus the selection of stimuli of test is determined during the experiment process, and stimulus order is decided by the pre-defined adaptive algorithm. Making relevant observations on the psychometric function with maximum efficiency without sacrificing accuracy is a common challenge of adaptive psychophysics.

With the success in many other fields, the theory of optimal experiments could be used as the basis for an adaptive psychometric procedure [7]. An optimal experiment requires a model of the process under control, and the parameters of the model can be estimated iteratively based on the outcomes of sequential previous presentations. This process includes two stages: parameters estimation and selecting next presentation.

Watson and Pelli (1983)[8] introduced a Bayesian framework to represent the probability distribution of the threshold parameter, but presentation levels were not optimized, which was improved by King-Smith et al. (1994)[9] to include optimal presentations. More recently, King Smith and Rose (1997) [10] developed a method specifically designed for measuring the slope of psychometric function; such function expresses the probability of detecting stimulus as a function of stimulus level [1]. In this method, each presentation is placed with the goal of maximizing efficiency. The method is adaptive as a stimulus level is decided by the probability density function generated from previous presentations, and the response at this level is used to update the probability of a given response like a function of real threshold, expressed as a likelihood function. Then the information provided by the likelihood function is combined with initial probability function by Bayesian multiplication. This procedure produces an updated function mapping the probability that each sound level is the threshold after the response to that stimulus. The updated probability function is then used to start next presentation and to make estimation of a threshold for optimum placement of the following presentation.

1.4. Gaussian Mixture Model

In the Bayesian optimal adaptive procedure, GMM model is built up with multi-Gaussian probability density function mixture, and is used for modeling the prior knowledge of psychometric experiments. A GMM model contains two parameters: the mean parameter displays the thresholds in the pure tone hearing test presentations and the variance parameter defines the spread of threshold values. The observed data-the Beltone database that contains over 400,000 audiograms, which were collected in U.S. Beltone offices over a 40-year period -is used as prior knowledge to estimate the parameters of GMM. During the testing procedure, the GMM parameters will be updated with new trial results again using recursive Bayesian procedure and maximum likelihood estimation.

1.5. Matlab based Graphical User Interface

Up till now, the basic idea for optimizing a psychoacoustic test procedure has been discussed. In the best case, when the system is working as expected at a high efficiency without sacrificing needed accuracy, we still have further problems to take into account: all the algorithms and methods are presented in theoretic and mathematical way, whereas the system is supposed to be used in a clinical environment where both the doctor and patient would prefer a more visible and direct manner to complete the test.

A graphical user interface is one sound choice meeting these requirements. It builds a close connection between user (doctor/patient) and mathematical modeling programs running on the computer and allows the tester to interact with computer programs in more ways other than typing. A fast correction or adjustment to the testing process can be easily accessed through this interface, while the response or opinion from the patient can also be taken into account by certain GUI functions.

But what kind of GUI are we going to build? In clinical cases, Audiometry testing is often done in two different ways: manually or automatically. Automatic testing method has its own advantages: it could be done in pre-defined order; it doesn't need much attention from tester; the test result will be saved in certain order. However under our PTA testing circumstance, flexible adjustment accessibility is required as the testing sound frequency/level is not under a pre-defined order but more of decision made by doctor based on his experience and system recommendation. What is more, for some rare cases like testing with too young child or aged people, it could be aborted midway which is unexpected but does happen. For the above reasons, a manually controlled GUI is chosen to fulfill the requirements of PTA tests.

The Matlab based GUI benefits the user more with its own outstanding calculation capacity and simple configurable interface construction. Adjustment, abortion and record loading/saving can be accessed comfortably through its interface in real time environment. Thus, a Matlab user interface can be expected to help improving the PTA testing procedure.

2. THEORY BACKGROUND

Several algorithms and methods are used in the implementation plan of GUI design for Bayesian optimal Pure Tone Audiometry, mainly concerned with pure tone Audiometry, Bayesian adaptive procedure, Gaussian Mixture Model simulation and Matlab GUI script.

2.1. Pure Tone Audiometry

As stated before, pure tone Audiometry (PTA) is a behavioral test used to diagnose hearing problems. Pure tone thresholds gained from this procedure indicate the softest sounds audible to the patient, and will be plotted on an audiogram to display sound level as a function of frequency. The extent of hearing loss is divided into six degrees [3]:

- Normal hearing (0-25 dB)
- Mild hearing loss (26-40 dB)
- Moderate hearing loss (41-55 dB)
- Moderate-severe hearing loss (56-70 dB)
- Severe hearing loss (71-90 dB)
- Profound hearing loss (>90 dB)

Before the test starts, the patient will be seated in a comfortable position. The head phones should be properly placed over the external auditory canal. This step should be carefully performed as the patient's pinna comes in various shapes and sizes, improper placement of head phones will cause threshold variations of even 15 - 20 dB. After the test started, the patient should respond in the way instructed before test, and he should do so even if the sound is only faintly heard.

The audiogram is obtained by measuring hearing thresholds at a frequency range from 125 Hz to 8 kHz by ascending method and an audiogram with both bone-conduction and air-conduction thresholds curve will indicate the type of hearing impairment for different individual [11]. Fig 2-1 shows an audiogram with mild gap between bone curve and air curve, which is inducted by conductive hearing loss. Fig 2-2 shows a high-frequency sensorineural hearing loss.

In ascending method, test sound starts at a very low level that it could not be heard by the patient, and then will increase gradually until the patient give a positive response to the test. Testing at each frequency is following the same rule, which means no matter what the previous threshold is, the test for new frequency has to be started all over from the lowest sound level. Thus the ascending method is argued to be inefficient and adaptive method is discussed to utilize the correlation between thresholds of adjacent frequencies.

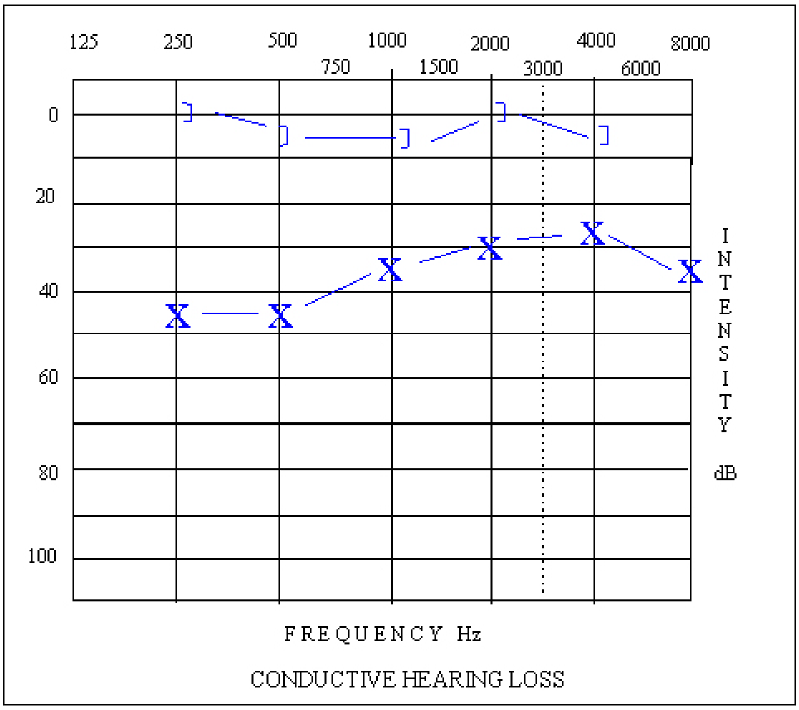

Fig 2-1 Audiogram indicating mild conductive hearing loss, square marked line is bone-conduction curve while cross marked line is air-conduction curve[11]

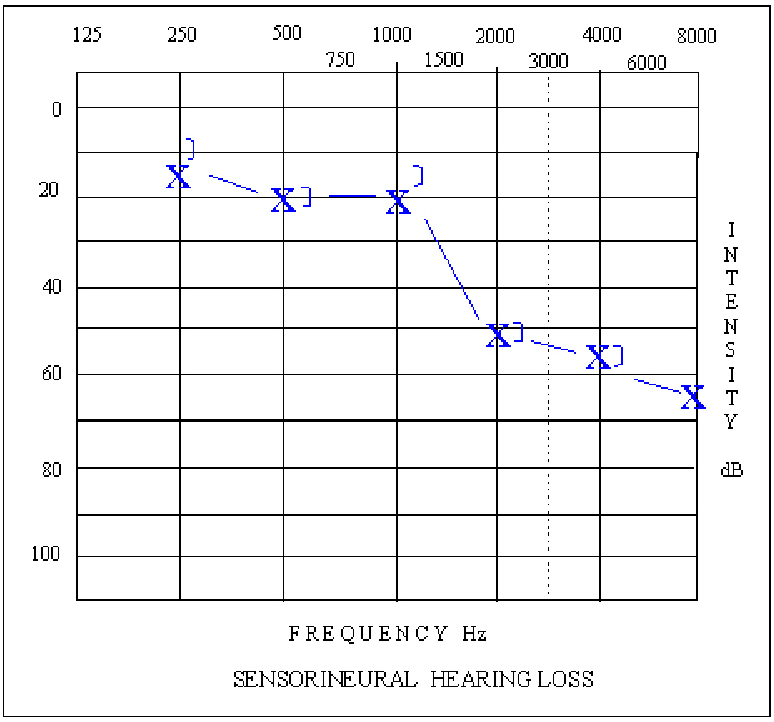

Fig 2-2 Audiogram indicating high-frequency sensorineural hearing loss, square marked line is bone-conduction curve while cross marked line is air-conduction curve[11]

2.2. Bayesian Optimal Algorithm

The inefficiency of ascending method makes developing a more rapid data collecting procedure strongly motivated. The idea to accomplish this goal is to use the theory of optimal experiments, in which Bayesian algorithm is used to include previous presentations' results. The Bayesian optimization is based on two steps:

Firstly, the parameters of the model are estimated to best fit the database-this part is done by GMM modeling and other related method like Vector quantization (VQ) etc.

Secondly, determining what stimulus level should be used for next presentation. After a new response R_k is received with presentation level L_k and frequency f_k at k time, parameter set θ is re-estimated with [1]:

$$f(\theta|R_k, L_k, f_k) = \frac{P(R_k|\theta, L_k, f_k)}{P(R_k)} f(\theta|R_{k-1}, L_{k-1}, f_{k-1}),$$

where f is the posterior probability distribution, $P(R_k|\theta, L_k, f_k)$ is the psychometric function. Here R_k is the kth response at level L_k and frequency f_k, and they also imply knowledge of previous presentations. R_k is a binary variable represents the patient's response Yes or No.

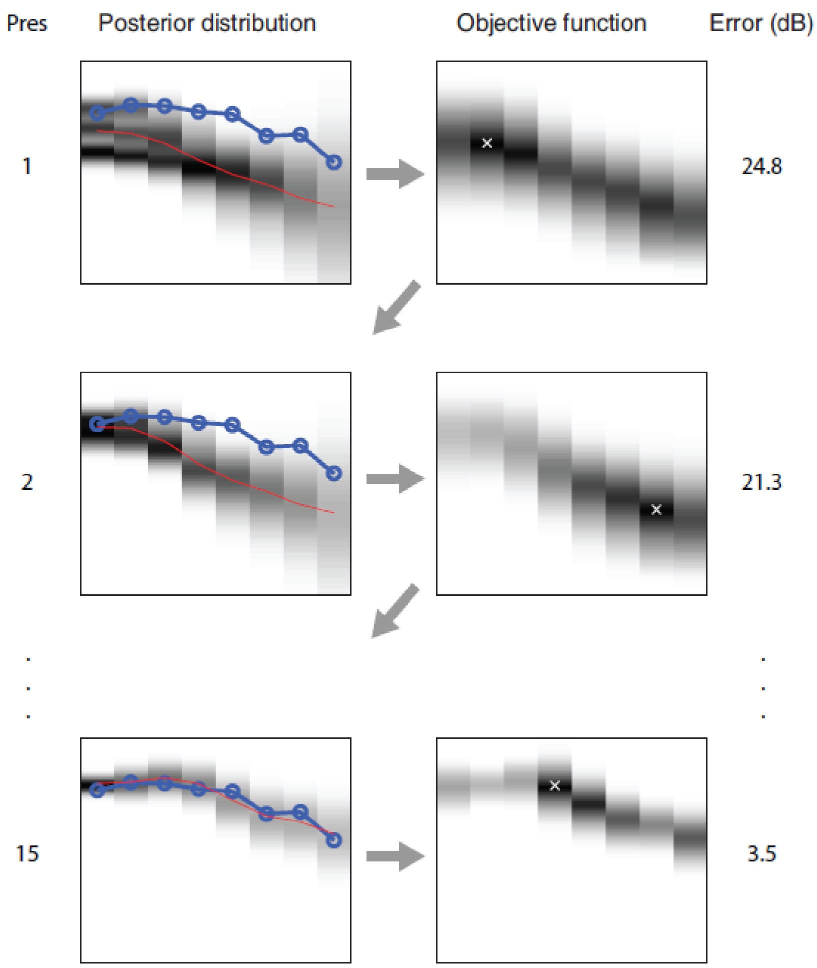

Fig 2-3 Recursive Bayesian Procedure for parameter estimation, blue circle marked line is a real hearing thresholds curve randomly chosen, red line is test curve, white cross mark is test point for each presentation, gray area indicates the probability real thresholds lie in such area[12]

In this method, the GMM with prior knowledge is updated through the testing and will be used to calculate the most probable hearing thresholds curve.

Figure 2-3 is an example for such procedure [12]. In the left column of the figure, the curve with blue round marks is one set of hearing thresholds randomly picked up from the Beltone database, and can be considered describing a real patient's hearing loss degree. The reason of displaying it here is that a comparison object is needed to identifying how large the gap is between the test result curve and the real thresholds curve. Thus, another red curve is plotted to display the test curve, which is initialized with a GMM. Test point is given as the right column of the figure shows, and the response will update the posterior probability distribution during the procedure and the estimated curve. While the iteration continues, the error between real thresholds curve and test result curve is decreasing significantly.

2.3. Gaussian Mixture Model

Gaussian Mixture Model (GMM) could almost smoothly simulate any shape of density distribution [13], thus it is used as the prior distribution for the hearing thresholds function's parameter. Using GMM as prior distribution is the key reason for the efficiency of Bayesian optimal method.

GMM is made up of N Gaussian distributions, when every single Gaussian distribution is called a component of the whole GMM. The integral GMM probability density which is linearly summed by these N Gaussian components can be explained as [14]:

$$g(x) = \sum_{i=1}^{N} g(i)g(x|i)$$

$$= \sum_{i=1}^{N} k_i \mathcal{N}(x|\mu_i, \Sigma_i)$$

where g(x) is the probability density function of event x. $g(i)$ is the probability function of *i*th component, and $g(x|i)$ is the probability of event x produced by *i*th component. This equation can be understood in another way, where k_i is the coefficient of *i*th component while k_i fulfils:

$$\sum_{i=1}^{N} k_i = 1,$$

and $\mathcal{N}(x|\mu_i, \Sigma_i)$ is the probability density function of every Gaussian component as it fulfils the same distribution as normal density.

The process of using GMM as the prior distribution of hearing thresholds parameter could be considered as a process of data clustering- the aiming data here is extracted from the Beltone database- or a training process. However, how do we use the database to 'train' the GMM? In more detail, we assume the database is generated by the GMM, after that we compute the density distribution function of GMM, so that the N components are corresponding to N clusters. This procedure is also called density estimation.

For now, the function is known to be GMM function, so the work left is to compute the parameters of GMM. Maximum likelihood Estimation and Expectation Maximization algorithm are used to calculate the parameters.

2.4. Maximum Likelihood Estimation

We are aiming to compute the parameters set to make GMM fit the observed data best, Maximum Likelihood Estimation (MLE) is what we are using to accomplish this work.

In general, when we are trying to obtain the value of parameters of a model which is used to describe a set of observations, MLE will help us estimate them by selecting the values most likely to produce the observations.

For our case, the density function is specified as Gaussian function (normal density distributed), and the parameters waiting to be estimated are k_i, μ_i, Σ_i. The basic idea is to find such parameter sets so that the probability density function has the largest possibility to produce the K observed data collectivity [14]. Such probability (likelihood) function is calculated by:

$$L(X) = \prod_{k=1}^{K} g(x_k)$$

$$= \prod_{k=1}^{K} \sum_{i=1}^{N} k_i \mathcal{N}(x_k | \mu_i, \Sigma_i)$$

X is the collectivity of events x_k, and we will get the logarithm of the equation above:

$$\log\big(L(X)\big) = \log \prod_{k=1}^{K} g(x_k)$$

$$= \sum_{k=1}^{K} \log \sum_{i=1}^{N} k_i \mathcal{N}(x_k | \mu_i, \Sigma_i)$$

Up to now, the idea to calculate the parameters set which give the function largest possibility to produce these K observed data has become calculating the parameters set to obtain maximum value of $\log(L(X))$. This step is done by Expectation Maximization.

2.5. Expectation Maximization

The parameter computation of GMM is accomplished by Expectation Maximization (EM), which is divided to two parts. First part is to estimate the probability that observed data is produced by every component of GMM. Second part is to estimate parameters for every component, assuming that observed data is produced by certain component of GMM. These two parts are iterated until the likelihood function becomes convergent.

There's an example shows the process of training new GMM [15]:

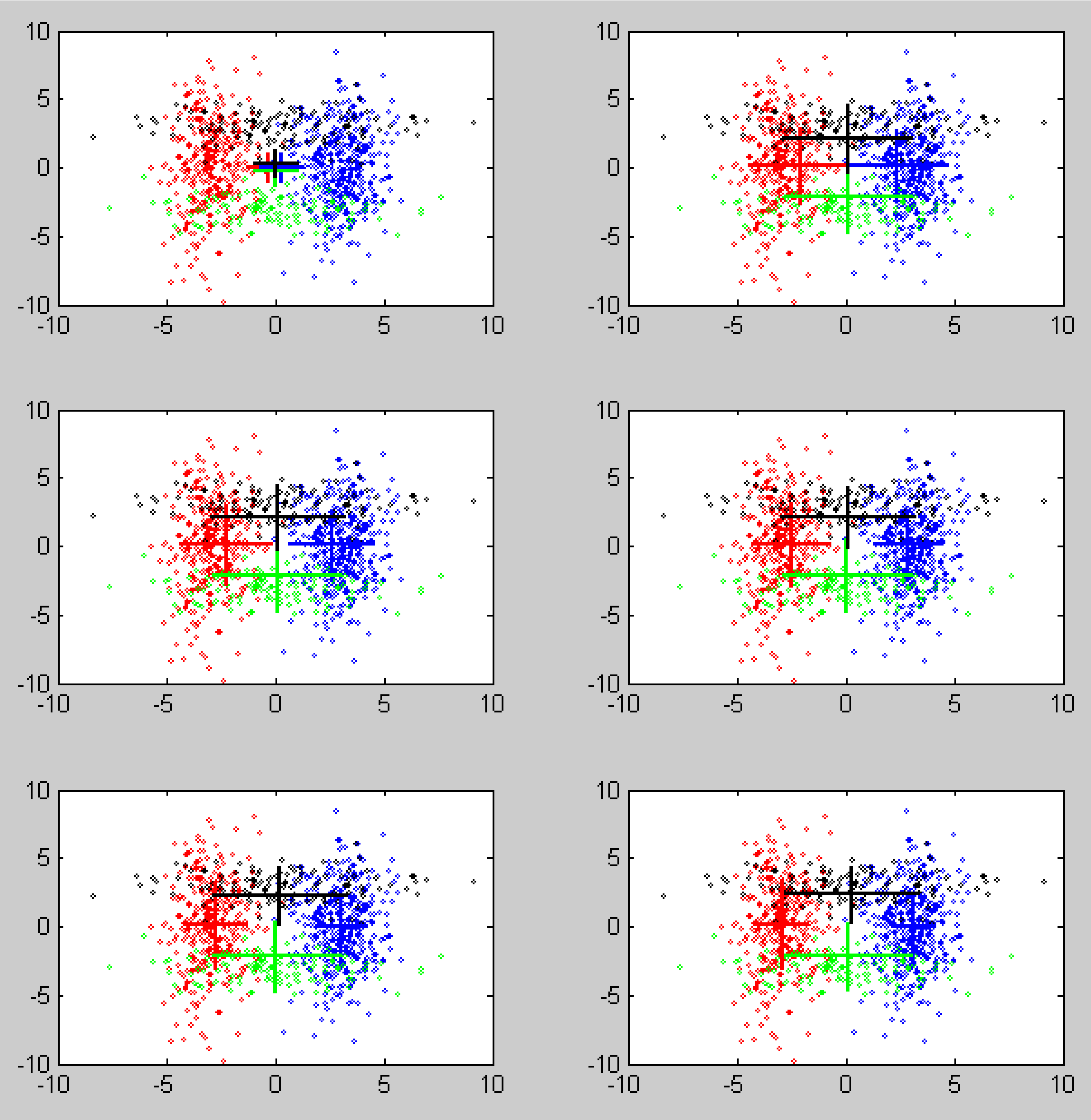

Fig 2-4 GMM training process example. Six plots are from left to right and top to down, four colors of dots are generated by an unknown GMM, and four crosses are generated by another GMM. During the training the second GMM is adapting to the dot distribution. [15]

In Fig 2-4, there're six plots here which arrange from left to right and top to down who shows the training process. The dots are produced by a GMM made up of four Gaussian models; while the color indicates the single Gaussian density function it's generated from. The other four big crosses represent a second GMM aiming to simulate the density distribution producing the dots. The dots distribution is unknown to the second GMM, after set up with initial parameters, the second GMM is adapting to the dots distribution gradually in the iteration related with MLE and EM algorithm.

From the plots we can see the four cross are approaching the middle of dots.

2.6. Graphical User Interface Expectation

A list of elements will be needed by a good GUI for testing and explaining the experiment result.

i. ***Audiogram Chart Unit***

Fig 2-5 below shows a standard audiogram:

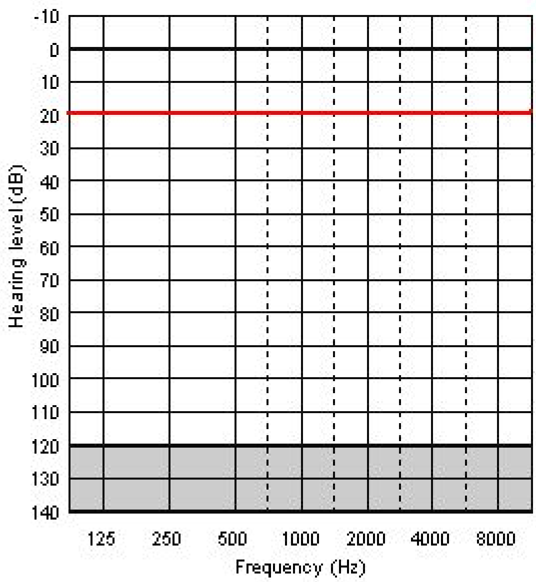

Fig 2-5 A standard audiogram example

As it shows, a chart with axis of frequency and hearing level, both ranges according to experiment requirement, is mandatory. This chart should be clickable; every clicked spots will be saved and translated to sound level and frequency messages, which will be used by audiogram to show the testing process and result.

ii. ***Input/Output Unit***

As a GUI, it will be convenient for the tester and patient to control and give response to the testing process, thus input and output signals should be easily accessible and adjustable.

Patient information input options, like age and gender will be needed for generating variant estimation curves for different patients. These curves were pre-estimated based on the GMM trained from extra database provided by GNResound.

What is more, button units should be built for patient to response to sounds stimuli. 'Yes'/'No' button will generate a positive/negative response to the estimation functions to update the estimated curves after each stimuli, 'Again' button indicates the patient need the sounds to be played again.

Output functions are mostly displayed in the audiogram chart, which includes the thresholds curve, testing points/results and gray scale density. Here the gray scale density histogram tells the tester the probability real thresholds lie in corresponding area. Such function is implemented by summing up the probability of normal distribution with related mean and variance parameters of each frequency/sound level set.

iii. *Control Unit*

Full control over the GUI is supposed to be given by control buttons group. After all the needed initial information is collected from the patient, 'start' command should be sent to the interface; when certain sound is needed to be played, or recommendation is needed by the doctor to make next presentation, corresponding instruction should be sent to the interface as well.

iv. *Feedback Unit*

During the testing process, parameters are supposed to be displayed visually. Thus a group of parameter text windows is used to supervise experiment parameters and operating conditions, including current clicking point position, estimation status and warning messages.

3. Implementation Method

This project is aiming to build a graphical user interface for Audiometry testing, and its main efficiencies are coming from the prior knowledge extracted from the Beltone database, which is the thresholds distribution of hearing impairment people in America, and the Matlab based user interface. The prior knowledge is represented by trained GMM and user interface is build up by Matlab UI class.

3.1. GMM Training

As the prior distribution of thresholds parameter, GMM is the basis of the system. The training process is based on Bayesian, Maximum Likelihood Estimation and Expectation Maximization algorithms which have been introduced before. In detail, it is done in the following steps:

i. *Database Extraction*

There is an earlier version of database which includes only hearing thresholds information of the patients. The corresponding GMM is extracted with thresholds only and saved in the work of Svante Stadler [1], who is the starter of this project.

The new extra data from Beltone database includes two parts of information, contact and hearing aid equipment information. Patient contact file covers ID, gender, birthday of the patients, the creation date of patient record file and the latest edit date of the record; hearing aid equipment file contains ID tag -which is identical to contact file, types of hearing aid instrument using by the patients and most important the hearing thresholds of the hearing aid instrument. Both files have large amount of redundant information which should be discarded, like record edited date, hearing instrument type etc. The rest of database will be used to generate and train the new GMM model.

This part of work is done in three Matlab functions: **merge_data.m**, **merge_contac.m**, **merge_wholedata.m**. Because there're around 400,000 portions of data in whole which is too large for Matlab to deal in one variable, both contact and equipment files are divided into five parts.

ii. *Erroneous Data Excluding*

After the extraction step, all the useful information is saved, while not all of them are available for training GMM. There are four cases for which corresponding data should also be discarded:

- Some thresholds values are missing and being replaced with '-1'.
- Some threshold values are not 5 dB quantized.
- In some rare cases, all threshold values of same patient are equal.

- Some threshold values are with suspicious leaps between adjacent frequencies, that is larger than 60 dB or even more. This case will be manually excluded.

The result of this step is saved in **audio_data.mat** (contains thresholds and index) and **audiogram.mat** (contains audiogram info for each person). The following plot (Fig 3-1) shows the rough age group population of thresholds data:

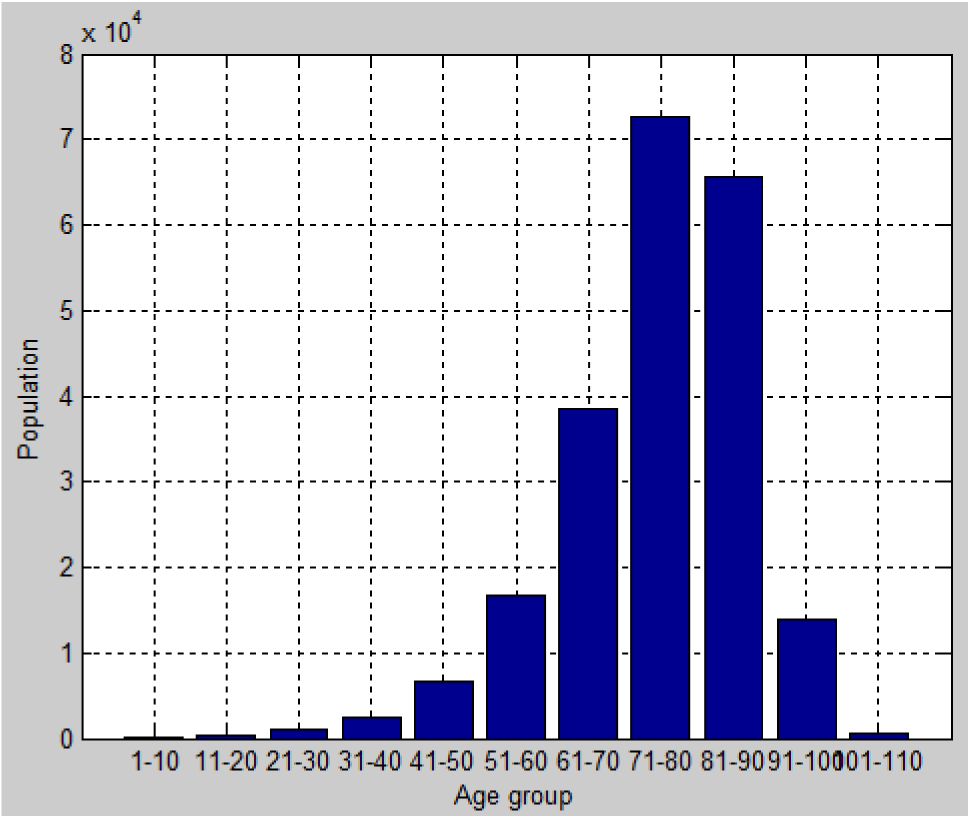

Fig 3-1 Population histogram as a function of age in Beltone database

In this plot, the horizontal axis represents age and vertical axis represents number of patients as a function of age.

iii. GMM Training With Extra Database

With valid information extracted from extra database, training can be started now. The number of components of GMM model should be large enough to simulate the distribution and avoiding overfitting the distribution to the valid data at the same time. As a reasonable compromise, ten components are chosen to build the GMM [14].

Firstly we set up a GMM including ten Gaussian distribution with same mean value: [60 60 60 60 60 60 60 60] and deviation value: [1 1 1 1 1 1 1 1]. Then the separated single Gaussian components are going to be initialized crudely by Vector quantization (VQ) clustering method with extracted data. After that, the data is used for the training process to refine the parameters of GMM, and such adaption process is based on MLE and EM algorithm which has been discussed in premier chapters.

Now the GMM is ready to be applied in the user interface as the prior knowledge of hearing thresholds. As the traditional PTA starts audio tests with lowest sounds level while Bayesian optimal method with prior knowledge starts tests at pre-estimated sounds level, it's expected to help improving the efficiency of Audiometry process.

3.2. Graphical User Interface

Matlab graphical user interface (GUI) has a short design period and outstanding calculation capability. Two methods are usually concerned while designing graphical user interface, one is using GUI guide provided by Matlab itself, in which a Visual Basic style toolbox is enabled to create different kinds of panels/handles/buttons; the other one is pure script coding in .m file, in which programmers describe all the frame elements one by one in Matlab code.

Firstly, the interface object is created like a class type, which means create a folder named @frame and save it with other clustering method classes. In this way the interface frame will be able to be reused by other similar interfaces. Secondly set, get and display functions are created for class @frame, which will enable designer to set up different parameters of the interface or display certain parameter value.

i. *GUI Style Choosing*

Several reasons make script coding the more viable method to design this interface:

- Children objects of the uitoolbox and axes are not fully parameterized implemented, part of the parameters can only be set up by script code.
- Uimenu cannot be created with editable features.
- Coded frames could be reused as standard template to create more interface elements.
- The entire user interface is being designed as a class type object, while Matlab GUI guide is not capable to generate such interface class.
- Positions and features of different control boxes are configured more accurately via script files, and the connection between input/output of panels and their callback functions are defined more clearly.

Concerning the interface is aiming for different control tools and configurable attributes, script code file style is used to build the interface.

ii. *GUI Elements*

- An audiogram chart in the interface is implemented like following:

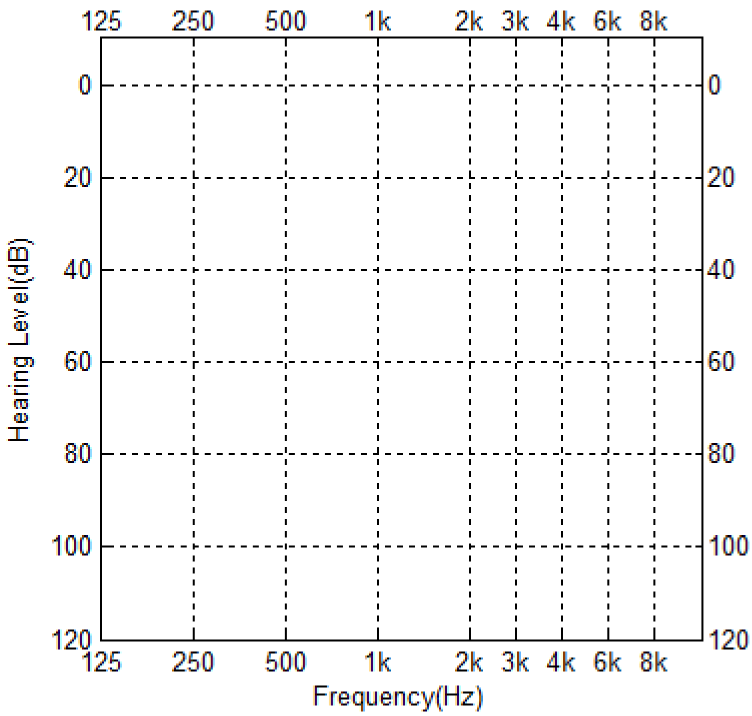

Fig 3-2 Audiogram chart frame used in GUI

In this chart, frequency ranges from 125 to 8k Hz and hearing level ranges from 0 to 120 dB.

- Input/output windows are used at the start of experiment, when gender and age data need to be selected, as the following figures show. Gender has three options, female/male/not specified, and age is divided by every 10 years. As soon as one of them is confirmed, an internal variable will save the choice and be used by the callback function to calculate the estimation curve. The estimation curve is calculated by applying **thresholdEstimate.m** function to the BayesMom psychometric estimator which is generated from age/gender parameters weighted GMM object.

Fig 3-3 Gender and Age input window in GUI

- Control buttons group is also embedded in the interface. 'Draw' button takes information from patient input, transfers them to estimation functions and draws the estimation curve extracted from database. 'Play' button plays the sound according to the position of clicked dots with different frequencies and sound levels. 'Recommend' button also needs input from experiment runner to indicate on which frequency a recommendation test point is needed, then the test frame will provide correspondent sound level.

Fig 3-4 Buttons for control usage, 'Draw' to drawing estimation curve and gray scale area, while 'Play' for playing test sound, 'Recommend' gives recommendation sound frequency and level for testing

- Testing process parameters are displayed in text window to give clear information about process operating conditions:

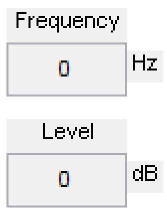

Fig 3-5 Parameters display window

- Warning/error messages are displayed in a separate procedure status window to represent the program running conditions. The status includes answer waiting, tone playing, recommendation thresholds estimating, re-estimating/updating curve and error warning etc.

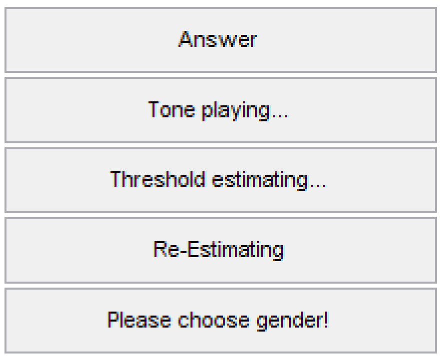

Fig 3-6 Warning/control message feedback window

3.3. Connecting Prior Knowledge (GMM) and GUI

Now we have gained the new GMM and a well established GUI, the next step is to bring them together to work as a whole. As gender and age are taken into account, they will decide the weight of GMM. After gender and age information is confirmed, the probabilities they are produced by every component are calculated and normalized to be used as the weight of GMM. In this way, different patients group are differentiated and simulated to estimate related hearing thresholds curve.

4. RESULTS

With the previous discussion, Bayesian optimization, GMM method and Matlab GUI design method are all used in this design; a following example will be represented to illustrate the result of implementation.

4.1. GMM with Extra Knowledge

The new GMM is trained by prior trials with extra parameters and maximizing the expected information gain on subsequent trials. The differential entropy function is created for the GMM weighted with gender and age information to make comparison with the GMM without gender or age information, and the following entropy result is obtained: in Fig 4-1, according to information theory, the entropy of gender and age information weighted GMM is lower than un-weighted GMM on all age groups, which means it is higher ordered and contains more available information.

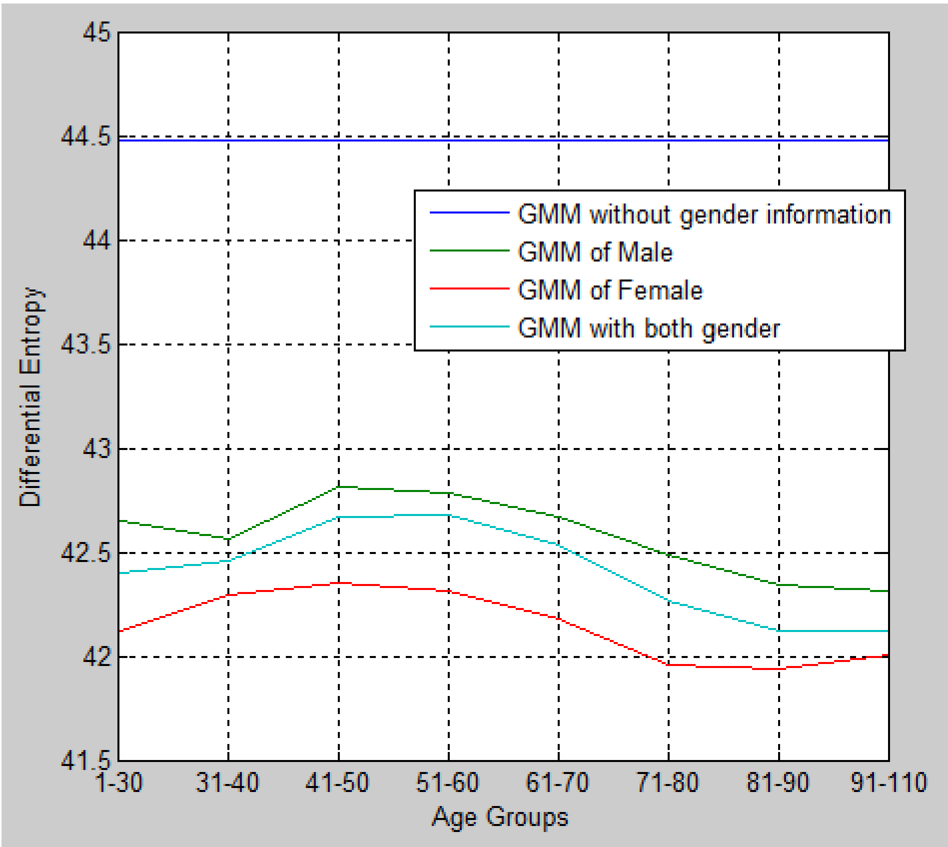

Fig 4-1 Entropy of trained GMM without gender or age information in dark blue line, entropy of GMM weighted with gender information on different age groups in green line (Male data only), red line (Female data only) and light blue line (data with both gender)

As a result, we can consider the new trained GMM with more parameters contains more information than the old one and can be more efficient when used to give tests to patient.

4.2. Audiometry Testing Initialization

The experiment starts with selecting age and gender data. Here we choose 'Female' for gender and '41-50' for age.

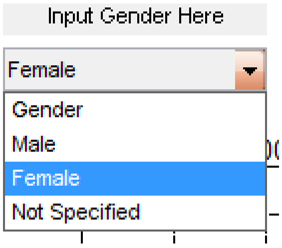

Fig 4-2 Gender input = Female

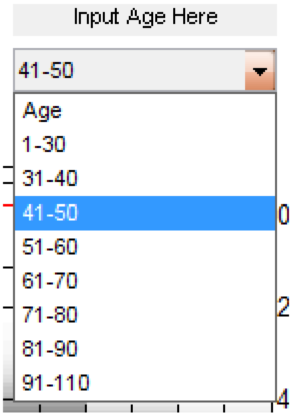

Fig 4-3 Age input = 41-50

After the selection is confirmed, both variables will be saved and a new GMM will be produced from the trained GMM with the weight related with given gender and age information. Then the estimation curve, which is supposed to be similar to current patient's hearing thresholds, will be calculated by estimating the mean parameter of the new GMM. 'Draw' button will send the instruction to plot it in the chart window. In this plot, estimation curve is showed in red line, and the gray area indicates the probability of hearing thresholds' existence in such area, the darker the area is, the higher probability it contains the hearing threshold.

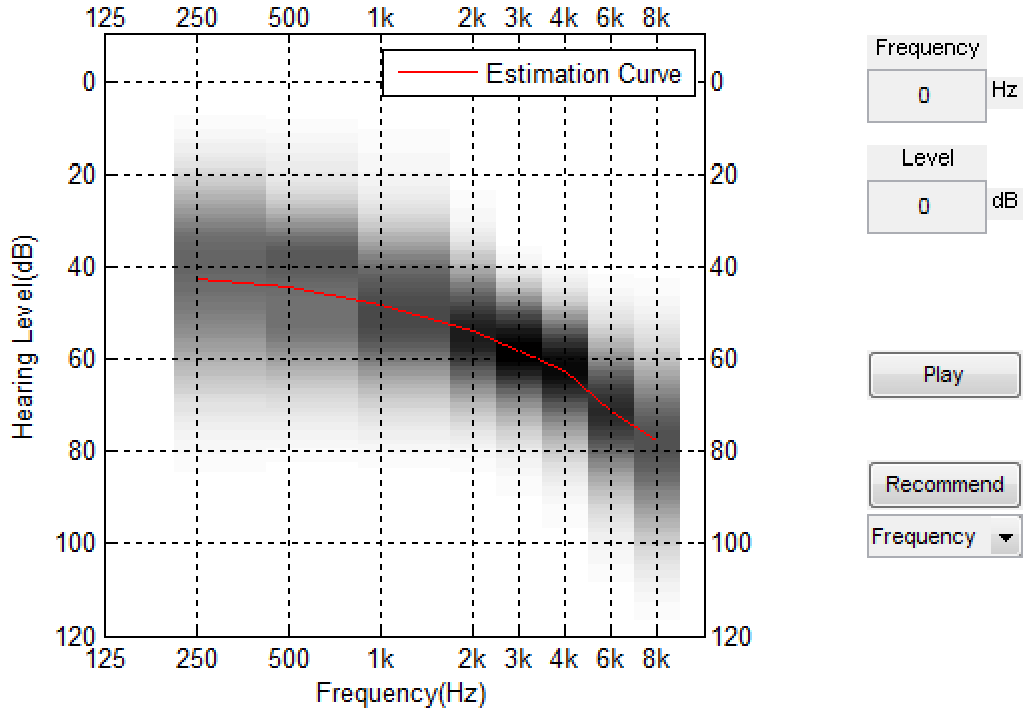

Fig 4-4 Initial estimation curve after age and gender information are confirmed, red line is estimation curve, gray area indicates the probability real thresholds lie in such area

The testing process can then be started from lowest frequency or any frequency which test runner prefer; also the runner could choose to start sound level manually (Fig 4-5) or by recommendation function (Fig 4-6).

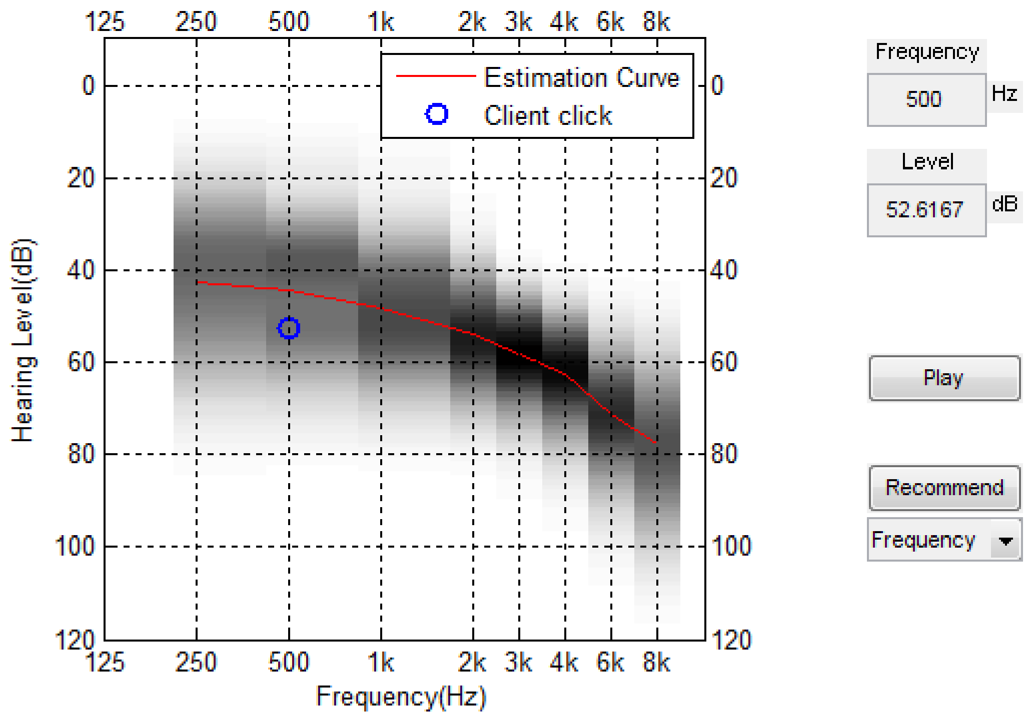

Fig 4-5 Tester's clicking point are displayed in blue circle, red line is estimation curve, gray area indicates the probability real thresholds lie in such area

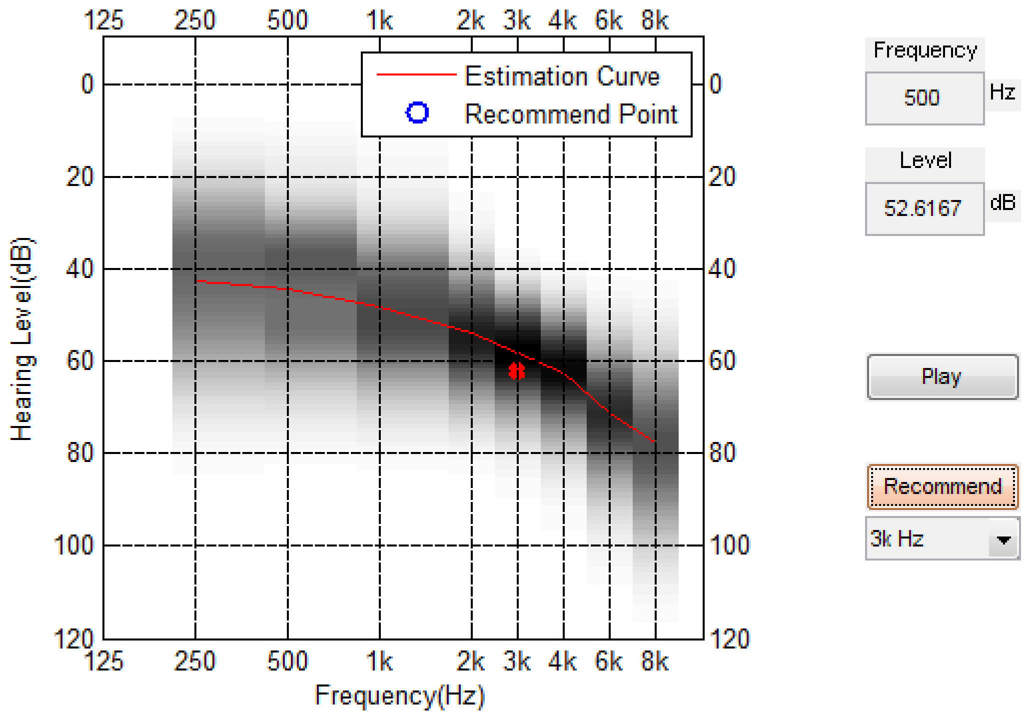

Fig 4-6 System recommendation point are displayed in red cross if no tester preferred point is chosen, red line is estimation curve, gray area indicates the probability real thresholds lie in such area

4.3. Hearing Thresholds Testing

After the test point is chosen and valid response is confirmed by the interface, the position of point will be displayed in frequency and sound level window. Every point tested and its response is saved in internal interface variables, to update the current estimation curve. Positive response and negative response will be saved separately and displayed in different color/type of dots in main chart, which are positive points in red circle dots and negative points in blue cross dots. In this way, a direct view of testing record and procedure is obtained by experiment runner.

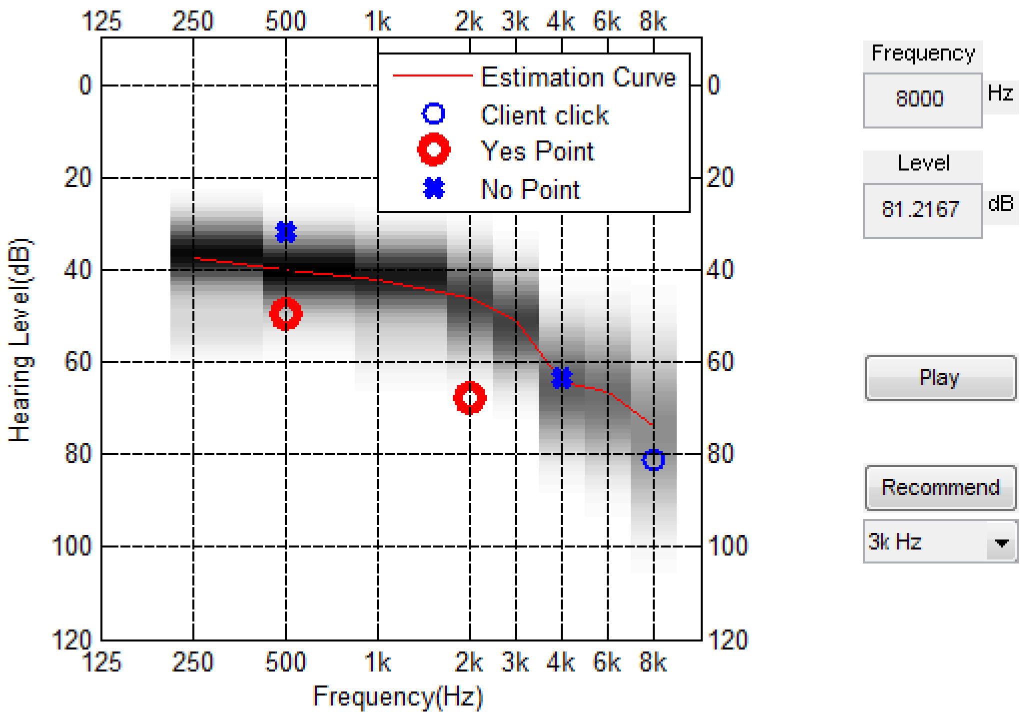

Fig 4-7 Example for test result after positive/negative responses are given. Red circles indicate positive answers are made while blue cross means the opposite. Blue circle stands for client clicking point

From the plot above, we can see the estimation curve is updated correctly due to response from patient, in which estimation curve is always beyond the positive response and under the negative response.

Back to our previous testing, in which gender is 'Female' and age is '41-50', suppose we have a patient is responding maturely to the tester. The chart starts like (Fig 4-8):

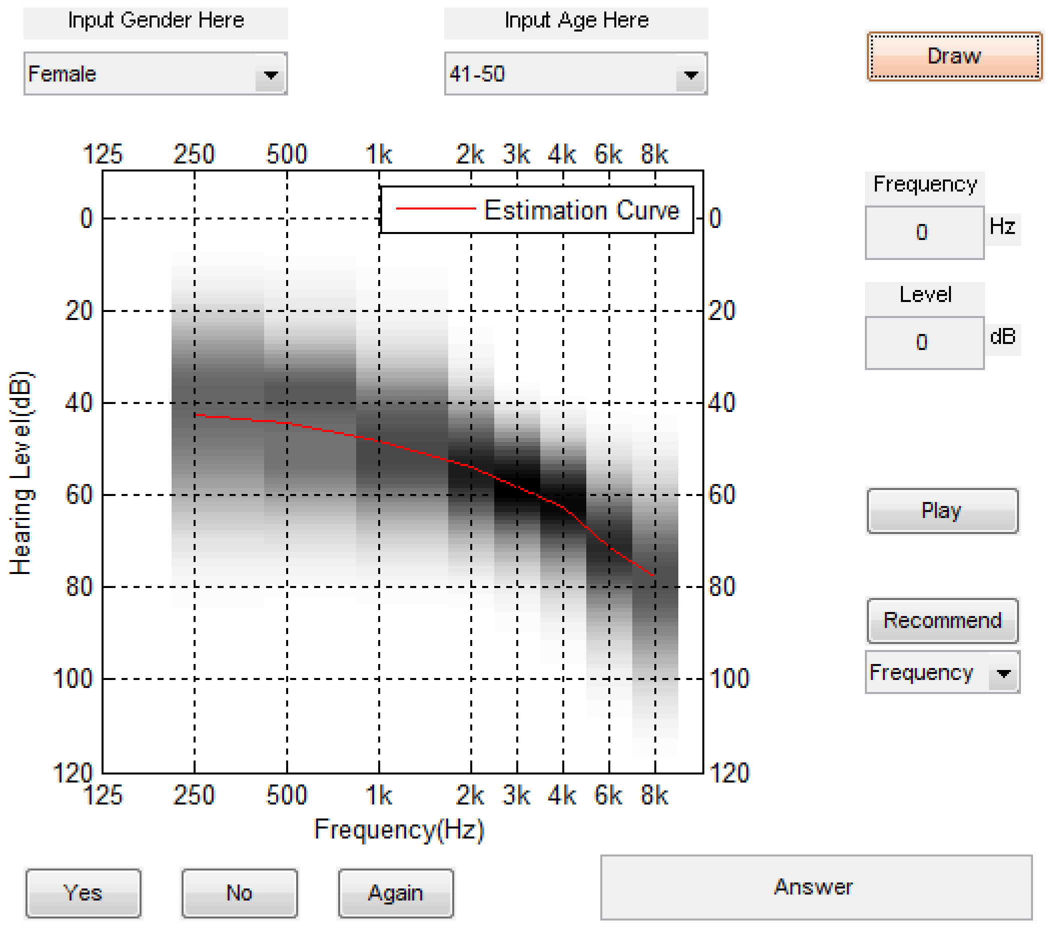

Fig 4-8 PTA test assuming a virtual patience attending, with gender = Female and age from 41 to 50, after Draw button pressed, an estimation curve is plotted.

Then click test points in the center chart from low levels to high levels, play the sound and wait for patient's response then decide the next test point. In this process, all points tested are saved and displayed.

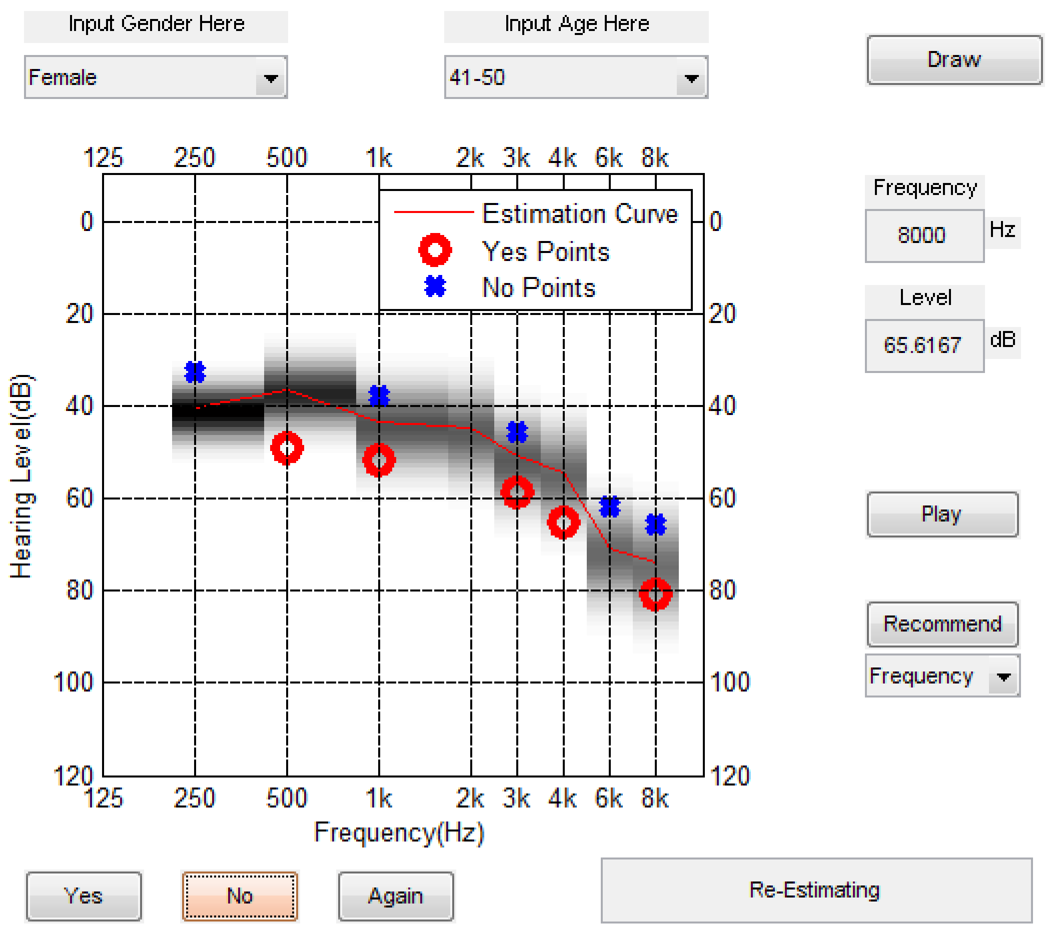

Fig 4-9 Test result after several presentations and responses are finished. All red circles mean positive answers while blue crosses the opposite.

All the red circles are positive response and all the blue crosses are negative response from patient. Meanwhile, the hearing thresholds curve is re-estimated and re-plotted.

5. DISCUSSION

5.1. Improvement Expected

In future studies for GUI of optimal pure tone Audiometry, more conditional data could be obtained like ear side. Normally, there is difference about hearing loss between two ears for individual, including such parameters in the model and make the distribution be extracted on these parameters to obtain more detailed audiograms. However, this would need a large amount of extra information and could distinctly slow down the initialization of new GMM. A more detailed highly featured model is expecting to be worth the extra work, yet it still needs to be proved.

The extent of hearing loss of different frequencies is not usually the same to each other, and the loss on certain frequency can largely affect the language understanding ability of patient. As a result, audiogram -that could prove on which frequency the hearing loss is taking severe effect on the patient's ability to understand- will be much more helpful on doctor's determination of treatment.

5.2. Future Plan

At the beginning of this thesis, this Matlab based user interface was supposed to be loaded on a real Audiometry testing instrument. Such connection programs could be written in MEX files and be called from .m files to execute C programs to drive the testing instruments. This work is suspended as no instruments can be applied from the company, and expected to be done in future if any machine available.

What is more, the testing procedure is done entirely by assuming a virtual patient attended, while no program/system can be proved to be valid and efficient unless a field test is run. Thus a real case field test will truly help improving this design as well.

6. CONCLUSION

Using GMM as prior knowledge with extra parameters (gender and age) for Pure Tone Audiometry test is implemented as intended. The use of prior knowledge efficiently improve the speed of estimation process of common types of hearing thresholds, while uncommon types estimation is slowed down simultaneously. The entropy of the new models also shows that it's working more efficiently than previous model without extra parameters. However, prior knowledge can be picked more carefully by adding more components or adding more filters to database selection to improve the performance of worst cases.

The Matlab based GUI is making good connection between control interface and optimized PTA process, it's making the test procedure to be more convenient and easy accessible. Instant adjustment or even midway abortion is taken care of so that the risk of running the test all over is avoided.

7. References

[1] Stadler, Svante, "Bayesian Optimal Pure Tone Audiometry with Prior Knowledge," *Sound and Image Processing Lab, KTH, Sc Lic Thesis*, 2009.

[2] Kochkin, S. "MarkeTrak VII: Hearing Loss Population Tops 31 Million People," *The Hearing Review*, vol. 12(7), pp. 16-29, July 2005.

[3] Martini A (Ed.), "European Group on genetics of hearing impairment," *Biomedical and Health Research Programme (HEAR) Infoletter 2*, November 1996, 8.

[4] Hear-it.org, More than 80 million Europeans with hearing loss. [Online].

Available: http://www.hear-it.org/page.dsp?area=857

[5] Wikipedia, Pure Tone Audiometry – PTA procedural standards. [Online].

Available: http://en.wikipedia.org/wiki/Pure_tone_audiometry

[6] Leek, R. Marjorie, "Adaptive procedures in psychophysical research," *Perception & Psychophysics*, vol. 63(8), pp. 1279–1292, 2001.

[7] Remus, J. Jeremiah and Collins, M. Leslie, "A comparison of adaptive psychometric procedures based on the theory of optimal experiments and Bayesian techniques: Implications for cochlear implant testing," *Perception & Psychophysics*, vol. 69(3), pp. 311–323, 2007.

[8] Watson, A. B. and Pelli, D. G, "QUEST: A Bayesian adaptive psychometric method," *Perception & Psychophysics*, 33(2):113–120, 1983

[9] King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., and Supowit, A, "Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation," *Vision Research*, 34(7):885–912, 1994.

[10] King-Smith, P. E. & Rose, D, "Principles of an adaptive method for measuring the slope of the psychometric function," *Vision Research*, 37, 1595-1604, 1997.

[11] Kutz, Joe Walter Jr. (2010, Mar.). Audiology, Pure-Tone Testing. [Online].

Available: http://emedicine.medscape.com/article/1822962-overview

[12] Test examples from Stadler, Svante's work about GMM training plotSystem in Matlab

[13] Jang, Jyh-Shing Roger. (2007) GMM: Gaussian Mixture models. [Online].

Available:
http://neural.cs.nthu.edu.tw/jang/books/dcpr/doc/08gmm.pdf

[14] Pluskid. (2009) Clustering (3): Gaussian Mixture Model. [Online]

Available: http://blog.pluskid.org/?p=39

[15] Test examples from Prof Leijon, Arne's work testGMM in Matlab